13.0 GLOSSARY[1]

Accuracy[2]: (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and one aspect of "relevance". Accuracy is highly dependent on the prevalence of positives in the population being examined.

Acute Toxic Class (ATC) method: An acute oral systemic toxicity test method based on testing groups of animals at fixed doses in a sequential manner. The lethality outcomes are used to classify a test substance into the appropriate GHS acute oral toxicity category.

Adjusted R^2: R^2 values that are adjusted for the relative proportion of data points to explanatory variables. Adjusted $R^2 = 1 - (1 - R^2)[(n - 1)/(n - k - 1)]$ where k = number of independent variables and n = number of observations. See "coefficient of determination."

ANOVA: One-way (and two-way) analysis of variance. ANOVA compares the measurements (continuous variables) of three or more groups when the data are categorized in one way (one-way) or two ways (two-way). ANOVA assumes that the populations compared are normally distributed and that the variances for the groups to be compared are approximately equal.

Assay[2]: The experimental system used. Often used interchangeably with "test" and "test method."

Biphasic dose-response: Dose-response in which cytotoxicity increases (as dose increases), plateaus, and then increases again. See **Section 2.6.3**.

[1] The definitions in this Glossary are restricted to their uses with respect to *in vitro* cytotoxicity testing and the NRU test methods.
[2] Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

Category prediction: The GHS hazard category that includes the predicted LD_{50} value for a test chemical.

Coded substances: Substances labeled by code rather than name so that they can be tested and evaluated without knowledge of their identity or anticipation of test results. Coded substances are used to avoid intentional or unintentional bias when evaluating laboratory or test method performance.

Coefficient of variation: A statistical representation of the precision of a test. It is expressed as a percentage and is calculated as follows:

$$\left(\frac{standard\ deviation}{mean}\right) \times 100\%$$

Coefficient of determination: In linear regression, it denotes the proportion of the variance in Y and X that is shared. Its value ranges between zero and one and it is commonly called called "R^2." For example, $R^2 = 0.45$, indicates that 45% of the variance in Y can be explained by the variation in X and that 45% of the variance in X can be explained by the variation in Y.

Concordance[2]: The proportion of all substances tested that are correctly classified as positive or negative. It is a measure of test method performance and one aspect of "relevance." The term is often used interchangeably with "accuracy." Concordance is highly dependent on the prevalence of positives in the population being examined. In the NICEATM/ECVAM study, concordance was used to describe the proportion of test substances that were correctly classified into GHS acute oral toxicity hazard categories, or to describe the proportion of test substances for which the laboratories obtained the same classification result.

Confluency: A state in which cells in culture come into contact with other cells in the same culture to form a complete sheet of cells (monolayer). For this study, confluency is determined as a percentage of cell coverage of the tissue culture vessel growth surface (e.g., cell monolayer has 80% confluency).

Cytotoxicity: The adverse effects resulting from interference with structures and/or processes essential for cell survival, proliferation, and/or function. For most chemicals, toxicity is a consequence of non-specific alterations in "basal cell functions" (i.e., via mitochondria, plasma membrane integrity, etc.), which may then lead to effects on organ-specific functions and/or death of the organism. These effects may involve the integrity of membranes and the cytoskeleton, cellular metabolism, the synthesis and degradation or release of cellular constituents or products, ion regulation, and cell division.

Definitive test: The main test of the cytotoxicity assay for determining the IC_{50}. The concentration closest to the range finder test IC_{50} serves as the midpoint of the concentrations tested in a definitive test. Compared to the range finder test, the definitive test uses a smaller dilution factor for the concentrations tested.

Discordant chemicals: Chemicals for which the LD_{50} is not accurately predicted by the IC_{50} (and the associated regression formula) or the GHS toxicity category is not accurately predicted by the IC_{50} (and the associated regression formula). Also referred to as “outliers.”

EDIT: Evaluation-guided Development of New *In vitro* Test Batteries. An international project coordinated by the Scandinavian Society for Cell Culture to develop new *in vitro* tests for toxicity and toxicokinetics to be incorporated into test batteries for predicting acute and chronic systemic toxicity.

Endpoint[2]: The biological process, response, or effect assessed by a test method.

Fixed Dose Procedure (FDP): An acute oral systemic toxicity test method based on testing groups of animals at fixed doses. Evident toxicity outcomes are used to classify a test substance into the appropriate GHS acute oral toxicity category.

F_G: An empirical factor for the RC regression line that represents the expected precision of LD_{50} predictions from basal cytotoxicity data. The LD_{50} values of 73% of the 347 RC chemicals are localized in the dose range around the RC regression line by $\mathbf{F_G} \leq \log 5$. The factor represents the expected difference between the LD_{50} determined in animal experiments and the LD_{50} estimated from the IC_{50} on the RC regression line.

Geometric mean: The antilog of the mean of the logarithm of the values. It is less affected by extreme values than the arithmetic mean.

Globally Harmonized System (GHS): A classification system presented by the United Nations that provides (a) a harmonized criteria for classifying substances and mixtures according to their health, environmental and physical hazards, and (b) a harmonized hazard communication elements, including requirements for labeling and safety data sheets.

Good Laboratory Practices (GLP)[2]: Regulations promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency, and principles and procedures adopted by the Organization for Economic Cooperation and Development and Japanese authorities that describe record keeping and quality assurance procedures for laboratory records that will be the basis for data submissions to national regulatory agencies.

Guidance Document: *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity* (ICCVAM 2001b).

Hazard[2]: The potential for an adverse health or ecological effect. A hazard potential results only if an exposure occurs that leads to the possibility of an adverse effect being manifested.

Hill function: The IC_{50} values are determined from the concentration-response using a Hill function which is a four parameter logistic mathematical model relating the concentration of the test chemical to the response (typically following a sigmoidal shape).

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\log IC50 - X)\text{HillSlope}}}$$

where Y= response, X is the logarithm of dose (or concentration), Bottom is the minimum response, Top is the maximum response, $\log IC_{50}$ is logarithm of X at the response midway between Top and Bottom, and HillSlope describes the steepness of the curve.

Hill function (revised): Some unusual dose-responses did not fit the Hill function well. To obtain a better model fit, the Bottom parameter was estimated without constraints (the previous practice was to use Bottom = 0). However, when Bottom ≠ 0, the EC_{50} reported by the Hill function was not the same as the IC_{50} since the Hill function relies on EC_{50} defined as the point midway between Top and Bottom. Thus, the Hill function calculation using the Prism® software was rearranged to calculate the concentration corresponding to the IC_{50} as follows.

$$X = \log EC_{50} - \frac{\log\left(\frac{Top - Bottom}{Y - Bottom} - 1\right)}{HillSlope}$$

X is the logarithm of concentration at 50% response, $\log EC_{50}$ is logarithm of concentration at the response midway between Top and Bottom, Top is the maximum response, Bottom is the minimum response, Y = 50 (i.e., 50% response), and HillSlope describes the steepness of the curve.

IC_{50}: test chemical concentration producing 50% inhibition of the endpoint measured (i.e., cell viability).

Interlaboratory reproducibility[2]: A measure of whether different qualified laboratories using the same protocol and test substances can produce qualitatively and quantitatively

similar results. Interlaboratory reproducibility is determined during the prevalidation and validation processes and indicates the extent to which a test method can be transferred successfully among laboratories.

Intralaboratory repeatability[2]: The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.

Intralaboratory reproducibility[2]: The first stage of validation; a determination of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times.

In vitro: In glass. Refers to assays that are carried out in an artificial system (e.g., in a test tube or petri dish) and typically use single-cell organisms, cultured cells, cell-free extracts, or purified cellular components.

In vivo: In the living organism. Refers to assays performed in multicellular organisms.

K_{ow}: Octanol:water partition coefficient.

LC_{50}: Acute lethal serum or blood concentrations.

LD_{50}: The calculated value of the oral dose that produces lethality in 50% of test animals (rats and mice). The LD_{50} values serve as reference values for the *in vitro* tests.

LD_{50} (initial): Acute oral rat and mouse LD_{50} values used during the chemical selection process. For RC chemicals, LD_{50} values were those used in the RC regression, which were largely from the 1983/84 RTECS®. For chemicals that were not included in the RC, the initial LD_{50} values came from HSDB or 2002 RTECS®.

LD_{50} (reference): Acute oral rodent LD_{50} values from rats and mice were located through literature searches and references from major toxicity databases such as RTECS®. Studies were reviewed to identify the most appropriate LD_{50} values for each chemical. Values obtained using feral animals, preanesthetized animals, or animals less than 4 weeks of age were not used. Values reported as inequalities were not used. Reference LD_{50} values were determined by calculating the geometric mean of the acceptable LD_{50} values. Data were used in generation of the 3T3 and NHK NRU regressions.

Maximum:minimum value: Ratio of minimum acceptable LD_{50} to maximum acceptable LD_{50}.

MEIC: Multicentre Evaluation of *In Vitro* Cytotoxicity. An international effort established by the Scandinavian Society for Cell Toxicology and initiated in 1983 to evaluate the relationship and relevance of *in vitro* cytotoxicity for predicting the acute toxicity of chemicals in humans.

Millimolar regressions: Linear regressions with IC_{50} values in mmol/L and LD_{50} values in mmol/kg.

Negative control: An untreated sample containing all components of a test system, except the test substance solvent, which is replaced with a known non-reactive material, such as water. This sample is processed with test substance-treated samples and other control samples to determine whether the solvent interacts with the test system.

Neutral red (NR): A weakly cationic water-soluble dye that stains living cells by readily diffusing through the plasma membrane and concentrating in lysosomes where it electrostatically binds to the anionic lysosomal matrix.

Neutral red uptake (NRU): Concentration of neutral red dye in the lysosomes of living cells. Altering the cell surface or the lysosomal membrane by a toxicological agent causes lysosomal fragility and other adverse changes that gradually become irreversible. The NRU

test method makes it possible to distinguish between viable, damaged, or dead cells because these changes result in decreased uptake and binding of NR measurable by optical density absorption readings in a spectrophotometer.

NHK: Normal **H**uman epidermal **K**eratinocytes (from neonatal foreskin).

Optical density (OD): The absorption (i.e., OD measurement) of the resulting colored solution (colorimetric endpoint) in the NRU assay measured at 540 nm ± 10 nm in a spectrophotometric microtiter plate reader using blanks as a reference

Outlier: For any measurement, an extreme value in the NICEATM/ECVAM study was referred to as an "outlier" if it passes a statistical test for outliers at the 99% level. With respect to chemicals, it refers to chemicals that do not fit (using the specified criteria) an IC_{50}-LD_{50} linear regression model. It may also refer to chemicals for which the predicted GHS toxicity category does not match the reference *in vivo* GHS toxicity category.

Performance[2]: The accuracy and reliability characteristics of a test method (see "accuracy", "reliability").

pH: A measure of the acidity or alkalinity of a solution. pH 7.0 is neutral; higher pHs are alkaline, lower pHs are acidic.

Plate reader: A spectrophotometric device for measuring light intensity as a function of color/wavelength (i.e., optical density/absorption at 540 nm ± 10 nm for NRU) in 96-well microtiter tissue culture plates.

Positive control: A sample containing all components of a test system and treated with a substance known to induce a positive response, which is processed with the test substance-treated and other control samples to demonstrate the sensitivity of each experiment and to allow for an assessment of variability in the conduct of the assay over time.

Predictivity[2]: Proportion of *in vivo* category matches for all substances with *in vitro* predictions for a particular category. Predictivity is an indicator of test accuracy.

Protocol[2]**:** The precise, step-by-step description of a test, including the listing of all necessary reagents, criteria and procedures for the evaluation of the test data.

Quality assurance (QA)[2]**:** A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures is assessed independently by individuals other than those performing the testing.

Range finder: Initial test performed to determine starting doses for the main (definitive) test. The NRU assays test eight concentrations of the test chemical or the positive control (PC) by diluting the stock solution in log dilutions to cover a large concentration range.

RC regression: $\log (LD_{50}) = 0.435 \times \log (IC_{50}) + 0.625$; for estimating an LD_{50} value in mmol/kg (body weight) from an IC_{50} value (in mM).

Reduction alternative[2]**:** A new or modified test method that reduces the number of animals required.

Reference substances: Substances selected for use during the research, development, prevalidation, and validation of a proposed test method because their response in the *in vivo* reference test method or the species of interest is known (see "reference test"). Reference substances should represent the classes of chemicals for which the proposed test method is expected to be used and cover the range of expected responses (negative, weak to strong positive).

Reference test method[2]**:** The accepted *in vivo* test method used for regulatory purposes to evaluate the potential of a test substance to be hazardous to the species of interest.

Refinement alternative[2]: A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal well-being.

Registry of Cytotoxicity (RC): Database that consists of *in vivo* acute oral toxicity data (i.e., LD_{50} values) from rats and mice and *in vitro* cytotoxicity data (i.e., IC_{50} values) from multiple cell lines and cytotoxicity endpoints for 347 chemicals with known molecular weights (Halle 1998). A regression model constructed from these data was proposed by ZEBET, as a method to reduce animal use by identifying the most appropriate starting doses for acute oral systemic toxicity tests

Relevance[2]: The extent to which a test method correctly predicts or measures the biological effect of interest in humans or another species of interest. Relevance incorporates consideration of the "accuracy" or "concordance" of a test method.

Reliability[2]: A measure of the degree to which a test method can be performed reproducibly within and among laboratories over time. It is assessed by calculating intra- and inter-laboratory reproducibility and intralaboratory repeatability.

Replacement alternative[2]: A new or modified test method that replaces animals with nonanimal systems or one animal species with a phylogenetically lower one (e.g., a mammal with an invertebrate).

Reproducibility[2]: The consistency of individual test results obtained in a single laboratory (intralaboratory reproducibility) or in different laboratories (interlaboratory reproducibility) using the same protocol and test substances (see intra- and inter-laboratory reproducibility).

RTECS®: Registry of Toxic Effects for Chemical Substances**.** Compendium of data extracted from the open scientific literature. The database includes toxicity data (e.g., acute toxicity) and specific numeric toxicity values (e.g., LD_{50}). Compiled by the U.S. National Institute for Occupational Safety and Health (NIOSH) and now licensed to MDL Information Systems, Inc.

Sensitivity[2]: The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy.

Simulation modeling: Computer simulation modeling of the acute systemic toxicity assays to determine animal use. The simulation process uses a simulated population of animals for testing, a reference endpoint (i.e., "true" LD_{50} value), and its assumed log-normal distribution. Morality is assumed to have a mean equal to the log of the true LD_{50}. The SD, which reflects the variability of the simulated population, is the inverse of the slope of the dose-mortality curve. Due to a lack of information for the real dose-mortality curve, the simulations assumed slopes of 0.5, 0.8, 2, 4, and 8.3.

Solubility: The amount of a test substance that can be dissolved (or thoroughly mixed with) culture medium or solvent. The solubility protocol was based on a U.S. EPA guideline (EPA 1998) that involves testing for solubility in a particular solvent, beginning at a relatively high concentration and proceeding to successively lower concentrations by adding more solvent as necessary for dissolution. Testing stops when, upon visual observation, the procedure produces a clear solution with no cloudiness or precipitate.

Solvent control: An untreated sample containing all components of a test system, including the solvent that is processed with the test substance-treated and other control samples to establish the baseline response for the samples treated with the test substance dissolved in the same solvent. When tested with a concurrent negative control, this sample also demonstrates whether the solvent interacts with the test system.

Specificity[2]: The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy.

Spirit of GLP: Guidance provided in the Statement of Work specifically for the non GLP-compliant laboratory that participated in the validation study. Based on the GLP standards referenced in the ECVAM Workshop 37 Report (Cooper-Hannan 1999) and the OECD Principles of GLP (OECD 1998). "*Laboratories that are non GLP-compliant shall adhere to*

GLP principles and other method parameters. Documentation and accountability shall be equal to GLP requirements. Laboratories must make assurances that they are equal in performance criteria and that there is parity amongst the laboratories."

TESS: Toxic Exposure Surveillance System. A comprehensive poisoning surveillance database maintained by the American Association of Poison Control Centers (AAPCC).

Test[2]: The experimental system used; used interchangeably with "test method" and "assay".

Test method[2]: A process or procedure used to obtain information on the characteristics of a substance or agent. Toxicological test methods generate information regarding the ability of a substance or agent to produce a specified biological effect under specified conditions. Used interchangeably with "test" and "assay". See also "validated test method" and "reference test".

Test method component: Structural, functional, and procedural elements of a test method that are used to develop the test method protocol. These components include unique characteristics of the test method, critical procedural details, and quality control measures.

3T3: BALB/c 3T3 clone A31 mouse fibroblasts developed in 1968 from disaggregated 14- to 17-day-old BALB/c mouse embryos (American Type Culture Collection [ATCC]; # CCL-163).

Tiered testing: A testing strategy where all existing information on a test substance is reviewed, in a specified order, before *in vivo* testing.

Toxicity underpredicted: Actual LD_{50} value of a test substance is lower than the predicted LD_{50} value.

Toxicity overpredicted: Actual LD_{50} value of a test substance is higher than the predicted LD_{50} value.

Transferability[2]: The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Up-and-Down Procedure (UDP): An acute oral systemic toxicity test method used to minimize the number of animals required to estimate the acute oral toxicity of a chemical, estimate the LD_{50} and confidence intervals (CI), and observe signs of toxicity. Single animals are tested sequentially. Subsequent doses are based on the outcome of the previous animal.

Validated test method[2]: An accepted test method for which validation studies have been completed to determine the accuracy and reliability of this method for a specific proposed use.

Validation[2]: The process by which the reliability and accuracy of a procedure are established for a specific purpose.

Vehicle control (VC): The VC consists of appropriate cell culture medium for the cells in the test (i.e., DMEM for 3T3 cells and keratinocyte growth medium for the NHK cells). For chemicals dissolved in DMSO, the VC consists of medium with the same amount of solvent as that used in the test chemical concentrations that are applied to the 96-well test plate. The final DMSO concentration is ≤ 0.5 % (v/v) in the VCs.

Volatility: Ability of a test chemical to evaporate. A general indicator of volatility issues in the NRU test methods is the percent difference in the mean OD values for the two VC columns on the test plate. If the difference is greater than 15%, then chemical volatility can be suspected, especially if the VC adjacent to the highest test concentration had a significantly reduced OD value. Volatility may be an issue for compounds with a specific gravity of less than 1.

Weight of evidence (process): The strengths and weaknesses of a collection of information are used as the basis for a conclusion that may not be evident from the individual data.

Weight regressions: Linear regressions with IC_{50} values in µg/mL and LD_{50} values in mg/kg.

ZEBET: The German National Center for the Documentation and Evaluation of Alternative Methods to Animal Experiments.